REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
30-10-2018	Final Report	27-Mar-2013 - 26-Mar-2018

4. TITLE AND SUBTITLE

Final Report: Information Aggregation for IED Identification with GPR, Video, and Electromagnetic Induction: Within Sensor Processing, Multi Sensor Fusion, and Large-Scale Learning

5a. CONTRACT NUMBER
W911NF-13-1-0065

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
633606

6. AUTHORS

5d. PROJECT NUMBER
622712

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES

Duke University
C/O Office of Research Support
2200 W. Main St., Ste. 710
Durham, NC 27705 -4677

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
62408-CS.20

12. DISTRIBUTION AVAILIBILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Leslie M. Collins
UU	UU	UU	UU		19b. TELEPHONE NUMBER 919-660-5260

RPPR Final Report
as of 24-Jan-2019

Agency Code:

Proposal Number: 62408CS **Agreement Number: W911NF-13-1-0065**
INVESTIGATOR(S):

Name: Leslie M. Collins Ph.D.
Email: lcollins@ee.duke.edu
Phone Number: 9196605260
Principal: Y

Name: Peter Torrione
Email: peter.torrione@duke.edu
Phone Number: 9196605252
Principal: N

Organization: **Duke University**
Address: C/O Office of Research Support, Durham, NC 277054677
Country: USA
DUNS Number: 044387793 EIN: 560532129

Report Date: 26-Jun-2018 Date Received: 30-Oct-2018
Final Report for Period Beginning 27-Mar-2013 and Ending 26-Mar-2018
Title: Information Aggregation for IED Identification with GPR, Video, and Electromagnetic Induction: Within Sensor Processing, Multi Sensor Fusion, and Large-Scale Learning
Begin Performance Period: 27-Mar-2013 **End Performance Period:** 26-Mar-2018
Report Term: 0-Other

Submitted By: Leslie M. Collins Email: lcollins@ee.duke.edu
Phone: (919) 660-5260

STEM Degrees: **STEM Participants:**

Major Goals: The fundamental objectives of this work are to use modern machine learning techniques to (1) develop new algorithms for both prescreening and object discrimination to support the HMDS program; (2) assess the utility of information-theoretic approaches that we have been developing for other sponsors for consideration by the various NVESD programs, particularly the incoming forward looking and handheld programs; (3) develop algorithms for any other sensors of interest to the sponsor and assess their performance. Historically, algorithm development work has included downward and forward looking GPR, IR, hyperspectral, acoustic, seismic, EMI, and video sensing modalities. For HMDS we have been carefully considering robustness issues with respect to target localization for feature extraction, modifying one of our previously developed prescreeners, and investigating convolutional neural networks as a new potentially effective processing algorithm. We work closely with other groups, including the system developer supported on these efforts, to insure both collaborative algorithm development (algorithm fusion) and technology transfer. In particular, we are actively leading an effort to generate a journal article submission that discusses the recent “bake off” blind test results (this journal paper has been returned with revisions). We also include some of our FLGPR work in this report.

Accomplishments: During the course of this project we have developed automated single sensor an fusion algorithms for handheld, HMDS, and FLGPR sensors. We have transitioned the algorithms that were developed to the sponsor and to the relevant contractors. We have also collaborated with others supported on these efforts to combine our algorithms, resulting in improved performance. In general, in tests involving all algorithm developers (both blind and otherwise) our algorithms performed at or above those of other developers.

Training Opportunities: 3 MS degrees were granted based on the support of this work, and and 7 Ph.D.s were granted.

Results Dissemination: Every year of the project we attended SPIE and presented 4-7 papers. Each graduate student published at least 1 paper, sometimes 2. A full list can be provided.

Honors and Awards: Nothing to Report

RPPR Final Report
as of 24-Jan-2019

Protocol Activity Status:

Technology Transfer: All code for algorithms of interest to the contractor/NVESD has been provided. Support for implementing this code (written in Matlab and C) was also provided.

PARTICIPANTS:

Participant Type: PD/PI
Participant: Leslie Marie Collins
Person Months Worked: 2.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Co PD/PI
Participant: Peter Acerbo Torrione
Person Months Worked: 12.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: Kenneth Dwayne Morton
Person Months Worked: 12.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Achut Manandhar
Person Months Worked: 12.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Nick Czarnek
Person Months Worked: 12.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)

RPPR Final Report
as of 24-Jan-2019

Participant: Daniel Reichmann
Person Months Worked: 12.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Rayn Sakaguchi
Person Months Worked: 12.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: Jordan Malof
Person Months Worked: 6.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Christopher Ratto
Person Months Worked: 12.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

ARTICLES:

RPPR Final Report
as of 24-Jan-2019

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: IEEE Transactions on Geoscience and Remote Sensing
Publication Identifier Type: DOI Publication Identifier: 10.1109/TGRS.2012.2215876
Volume: 51 Issue: 5 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Target Classification and Identification Using Sparse Model Representations of Frequency-Domain Electromagnetic Induction Sensor Data
Authors:
Keywords: EMI, Wideband, sparse, bayesian
Abstract: Frequency-domain electromagnetic induction (EMI) sensors can measure object-specific signatures that can be used to discriminate landmines from harmless clutter. In a model-based signal processing paradigm, the object signatures can often be decomposed into a weighted sum of parameterized basis functions, such as the discrete spectrum of relaxation frequencies (DSRF), where the basis functions are intrinsic to the object under consideration and the associated weights are a function of the target-sensor orientation. The basis function parameters can then be used as features for classifying the target. One of the challenges associated with effectively utilizing a model-based signal processing paradigm such as this is determining the correct model order for the measured data, as the number of basis functions containing fundamental information regarding the target under consideration is not known a priori. In this paper, sparse Bayesian relevance vector machine (RVM) regression is applied
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: IEEE Transactions on Geoscience and Remote Sensing
Publication Identifier Type: DOI Publication Identifier: 10.1109/TGRS.2013.2252016
Volume: 0 Issue: 0 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Histograms of Oriented Gradients for Landmine Detection in Ground-Penetrating Radar Data
Authors:
Keywords: HOG, computer vision, GPR, landmine, improvised explosive
Abstract: Detection of buried explosive threats is a challenging problem. GPR has recently become a powerful tool for achieving robust subsurface target detection, but novel target types, and large numbers of subsurface objects in rural environments significantly complicate accurate discrimination of explosive threats from harmless false alarms. Significant research in feature extraction from GPR data has previously shown the capability for improved performance. Similarly, many techniques from the computer vision literature have made significant strides in recent years in for improvements in object class recognition. This work studies the relationships between and application of feature descriptor techniques from the computer vision community in application to target detection in GPR data. Relationships between a very successful computer vision technique (Histogram of Oriented Gradients) and a related powerful technique from subsurface sensing (Edge Histogram Descriptors) are explored, and preli
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

RPPR Final Report
as of 24-Jan-2019

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: IEEE Transactions on Geoscience and Remote Sensing
Publication Identifier Type: DOI Publication Identifier: 10.1109/TGRS.2013.2257175
Volume: 0 Issue: 0 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Bayesian Context-Dependent Learning for Anomaly Classification in Hyperspectral Imagery
Authors:
Keywords: Context dependent, bayesian, hyperspectral, remote sensing
Abstract: Many remote sensing applications involve the classification of anomalous responses as either objects of interest or clutter. This paper addresses the problem of anomaly classification in hyperspectral imagery (HSI) and focuses on robustly detecting disturbed earth in the long-wave infrared (LWIR) spectrum. Although disturbed earth yields a distinct LWIR signature that distinguishes it from the background, its distribution relative to clutter may vary over different environmental contexts. In this paper, a generic Bayesian framework is proposed for training context-dependent classification rules from wide-area airborne LWIR imagery. The proposed framework combines sparse classification models with either supervised or discriminative context identification to pool information across contexts and improve classification overall. Experiments are performed with data from a LWIR landmine detection system. Contexts are learned from endmember abundances extracted from the background near each d
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal:
Publication Identifier Type: DOI Publication Identifier: 10.1117/12.2016063
Volume: 0 Issue: 0 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Sparse model inversion and processing of spatial frequency-domain electromagnetic induction sensor array data for improved landmine discrimination
Authors:
Keywords: bayesian, sparse, model inversion, EMI, RVM
Abstract: Frequency-domain electromagnetic induction (EMI) sensors have been shown to provide target signatures which enable discrimination of landmines from harmless clutter. In particular, frequency-domain EMI sensors are well-suited for target characterization by inverting a physics-based signal model. In many model-based signal processing paradigms, the target signatures can be decomposed into a weighted sum of parameterized basis functions, where the basis functions are intrinsic to the target under consideration and the associated weights are a function of the target sensor orientation. When sensor array data is available, the spatial diversity of the measured signals may provide more information for estimating the basis function parameters. After model inversion, the basis function parameters can form the foundation of model-based classification of the target as landmine or clutter. In this work, sparse model inversion of spatial frequency-domain EMI sensor array data followed by target c
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

RPPR Final Report
as of 24-Jan-2019

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal:
Publication Identifier Type: DOI Publication Identifier: 10.1117/12.2016113
Volume: 0 Issue: 0 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: A novel framework for processing forward looking infrared imagery with application to buried threat detection
Authors:
Keywords: forward looking, GPR, IR, landmines, buried threats, RX
Abstract: Forward Looking Infrared (FLIR) cameras have recently been studied as a sensing modality for use in buried threat detection systems. FLIR-based detection systems benefit from larger standoff distances and faster rates of advance than other sensing modalities, but they also present significant signal processing challenges. FLIR imagery typically yields multiple looks at each surface area, each of which is obtained from a different relative camera pose and position. This multi-look imagery can be exploited for improved performance, however open questions remain as to the best ways to process and fuse such data. Further, the utility of each look in the multi-look imagery is also unclear: How many looks are needed, from what poses, etc? In this work we propose a general framework for processing FLIR imagery wherein FLIR imagery is partitioned according to the particular relative camera pose from which it was collected. Each partition is then projected into a common spatial coordinate syste
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: IEEE Transactions on Geoscience and Remote Sensing
Publication Identifier Type: DOI Publication Identifier: 10.1109/TGRS.2014.2346954
Volume: 53 Issue: 4 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Multiple-Instance Hidden Markov Model for GPR-Based Landmine Detection
Authors:
Keywords: GPR, HMM, landmine detection, MIL
Abstract: Hidden Markov models (HMMs) have previously been successfully applied to subsurface threat detection using ground penetrating radar (GPR) data. However, parameter estimation in most HMM-based landmine detection approaches is difficult since object locations are typically well known for the 2-D coordinates on the Earth's surface but are not well known for object depths underneath the ground/time of arrival in a GPR A-scan. As a result, in a standard expectation maximization HMM(EM-HMM), all depths corresponding to a particular alarm location may be labeled as target sequences although the characteristics of data from different depths are substantially different. In this paper, an alternate HMM approach is developed using a multiple-instance learning (MIL) framework that considers an unordered set of HMM sequences at a particular alarm location, where the set of sequences is defined as positive if at least one of the sequences is a target sequence; otherwise, the set is defined as negati
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

RPPR Final Report
as of 24-Jan-2019

Final Report:
Information Aggregation for IED Detection With GPR, Video and Electomagnetic Induction: Within Sensor Processing, Multi-sensor Fusion, and Large Scale Learning

Grant Number 62408-CS

Leslie M. Collins, Jordan Malof, Daniel Reichman
Electrical and Computer Engineering Department
Duke University

Abstract

The fundamental objectives of this work are to use modern machine learning techniques to (1) develop new algorithms for both prescreening and object discrimination to support the HMDS program; (2) assess the utility of information-theoretic approaches that we have been developing for other sponsors for consideration by the various NVESD programs, particularly the incoming forward looking and handheld programs; (3) develop algorithms for any other sensors of interest to the sponsor and assess their performance. Historically, algorithm development work has included downward and forward looking GPR, IR, hyperspectral, acoustic, seismic, EMI, and video sensing modalities. For HMDS we have been carefully considering robustness issues with respect to target localization for feature extraction, modifying one of our previously developed prescreeners, and investigating convolutional neural networks as a new potentially effective processing algorithm. We work closely with other groups, including the system developer supported on these efforts, to insure both collaborative algorithm development (algorithm fusion) and technology transfer. In particular, we are actively leading an effort to generate a journal article submission that discusses the recent "bake off" blind test results (this journal paper has been returned with revisions). We also include some of our FLGPR work in this report.

Technical Progress

We present several areas of technical progress that were achieved over the course of this project and one area of related research that is relevant to a program being transitioned to a sponsor. We have previously transitioned a prescreener called HOG for the HMDS system to the government and the sensor developer. We have also continued to transition our algorithms that we developed for localizing the areas of interest in GPR imagery, and have worked with the other developers to share code and output files. We also have investigated a new algorithm to the area of landmine detection – convolutional neural nets. Algorithmic development efforts are described below. We also describe some related work in forward looking radar since that work will be transitioning to the sponsor.

A new feature descriptor, and discrimination algorithm, for downward-looking ground penetrating radar (GPR)

A large body of published research has focused on developing algorithms that automatically detect buried threats in GPR data. Most contemporary algorithms consist of two steps: a feature extractor, and a machine learning classifier. The feature extraction step attempts to make a succinct encoding of the original GPR data that retains all information relevant to discriminating between threats and non-threats, while suppressing any other variations in the data (i.e., noise). Machine learning algorithms are then trained to discriminate between threats and non-threats based upon the feature values computed on the GPR data (and corresponding spatial location) under consideration. Machine learning is a mature area of research, which has produced a large number of effective learning algorithms. Therefore, much of recent research for GPR buried threat detection (BTD) has been focused on the development of more effective features, which can result in substantial performance improvements.

Many existing GPR BTD algorithms operate on relatively small images extracted from larger GPR volumes, or B-scans (i.e., large images formed from GPR data). An example of such an image is shown below. Two particularly successful recent examples of GPR features operating on this type of imagery are Histogram of Oriented gradients (HOG) and the Edge Histogram Descriptors (EHD). The motivation for both HOG and EHD is to encode shape information within the GPR imagery, such as the dominant orientation of image gradients within smaller sub-regions, often called "pooling regions" within the GPR imagery. The pooling regions used for EHD, as applied to a GPR image, are illustrated below.

In our recent work we analyzed the weaknesses of the EHD and HOG features. Our analysis is *summarized* in images (b)-(d) below. These images are each an EHD feature obtained by averaging the values of EHD features extracted across three populations of GPR imagery, respectively: easy threats, difficulty threats, and non-threats. Easy buried threats are those that EHD-based algorithms can identify without making many, or any, false detections. Any improvements in detection performance therefore are unlikely to come from increased detection performance on these threats. Therefore, we focus our analysis on the identification of difficult threats, which are those that are frequently confused with non-threats. Notice that the vertical gradients (top row) within the Non-threat (b) and the Difficult threat (d) EHD feature averages have a high magnitude, and are very similar. Ideally our feature descriptor would suppress image content which causes threats and non-threats to appear similarly, and emphasize content that differentiates them.

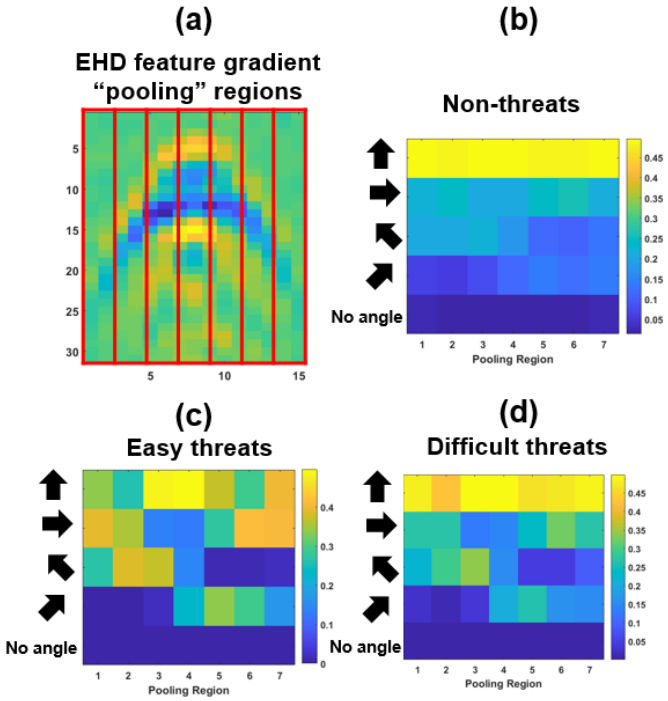

Figure 1: (a) Illustration of the pooling regions (tall red rectangles) used for the EHD feature. A pooling region is a region over which shape information is summarized, often with a gradient histogram, in the source image. (b)-(d) are illustrate the form of the resulting feature descriptor. There are five different shape categories, which vary depending upon their angle, unless there is no dominant angle. (b)-(d) are the EHD feature vectors obtained by averaging the feature vectors extracted over non-threats, easy threats, and difficult threats, respectively.

Based on our findings with EHD through numerous experiments (which are summarized below) we hypothesized that we might obtain better detection performance by suppressing the vertical component of EHD (and HOG) features. This hypothesis formed the basis for a new feature we recently developed, referred to thus far as the "T-scan" feature. The quality of the T-scan feature that differentiates it from features such as HOG and EHD is illustrated in Figure 2. The main idea of T-scan is to focus on extracting horizontal gradient components from the GPR imagery rather than vertical components. We hypothesized that suppressing this characteristic of these feature representations may improve detection performance, by emphasizing the differences in the shape content within threat and (difficult) non-threat GPR imagery, respectively.

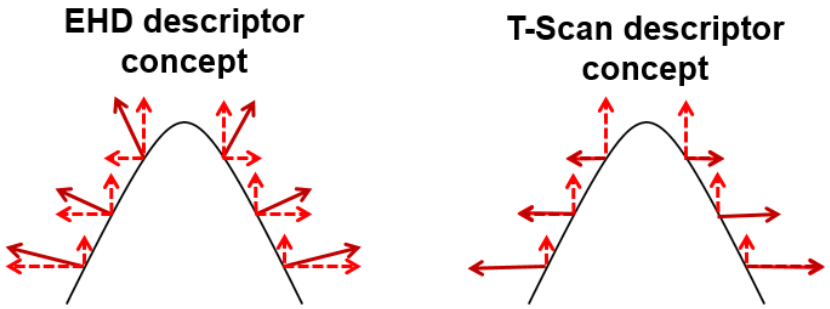

Figure 2: Illustration of the conceptual difference between the EHD descriptor (or feature), and the T-Scan descriptor. The EHD descriptor (left) attempts to encode the shape information given by the dark red arrows. T-Scan attempts to suppress this vertical shape information , and retain the horizontal shape component from the object in the GPR imagery. A hyperbolic shape is shown here for illustration, but in real-world GPR data, difficult threats do not always exhibit this canonical hyperbolic shape.

We leveraged this idea in the development of the T-scan feature, which extracts shape information from GPR imagery constructed from GPR data collected at a fixed time. We refer to this type of image as a T-scan, and its vertical and horizontal axes are both spatial. The T-scan relies on identifying and encoding shape information in T-scans, and thereby vertical shape content of the GPR data is ignored; only horizontally oriented shape information is encoded. These shape features are extracted over many time slices, and subsequently averaged, in order to create a final T-scan feature. This averaging results in a compact encoding of the total shape content within a volume of GPR data. This feature extraction process is illustrated in Figure 3.

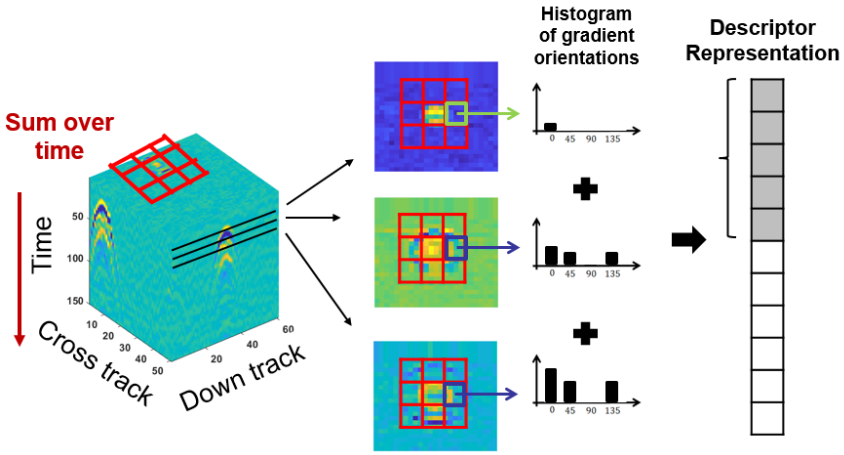

Figure 3: Illustration of the T-Scan feature computation on a volume of GPR data (left). Shape information is encoded in several T-scans extracted from the GPR volume (middle). The shape encoding takes the form of a gradient histogram, and the histograms from each T-scan are averaged to obtain a final feature, or descriptor (right).

In order to test the validity of the T-scan feature, and thereby test our feature extraction hypothesis, we compared our T-scan algorithm (i.e., T-Scan feature with a support vector machine classifier) with several other recently developed algorithms proposed by our university collaborators. This comparison was conducted as part of a process of evaluating algorithms for inclusion on the fielded GPR BTD system, and was moderated by the sponsoring agency (countermine). The performance of the algorithms was compared using lane-based cross-validation on a dataset that included 120,000 m^2 of inspected pathway, and 4552 threat encounters. The results of this performance comparison on shown in Figure 4. The T-scan algorithm obtains the best overall performance (as measured by the area under the ROC curve). The T-scan algorithm also ranks well (first or second) in performance regardless of the operating point. In those cases where it ranks second in performance, it obtains similar performance to the best algorithm.

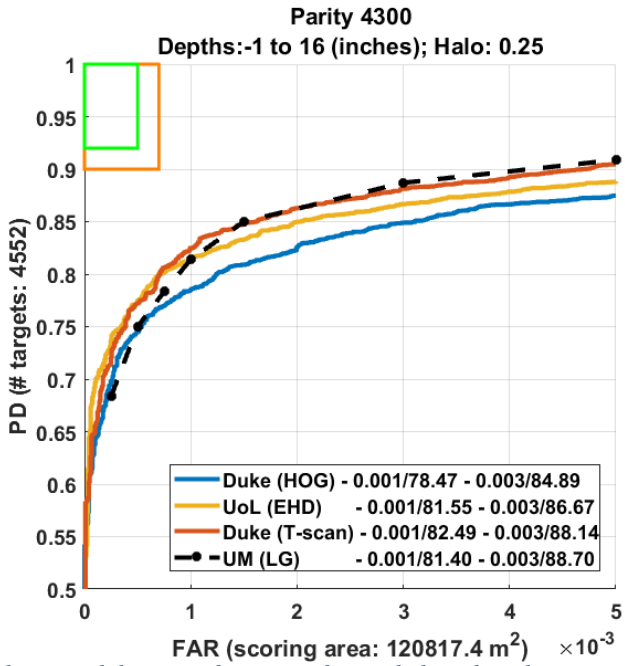

Figure 4: Lane-based cross-validation performance of several algorithms that were recently submitted for evaluation by the sponsoring agency.

Convolutional neural networks for recognition of buried threats in ground penetrating radar (GPR) data

As discussed, a large body of published research has focused on developing algorithms that automatically detect buried threats in GPR data. Many of these algorithms operate on small images extracted from larger GPR volumes, or B-scans. Perhaps as a result of this, many recent GPR detection algorithms have adopted techniques for image recognition from the computer vision literature. Some examples include Histogram of Oriented gradients (HOG), Edge Histogram Descriptors (EHD), and Fourier features.

Recently, deep convolutional neural networks (CNNs) have achieved impressive performance for image recognition tasks on natural images. This suggests that CNNs may also yield improvements for threat recognition in GPR data. A CNN is comprised of a set of learned feature extractors connected with a neural network classifier. This is unlike traditional object detection pipelines which typically separate the extraction of a static set of features (e.g., HOG or EHD) and the classification of the data. The CNN is comprised of layers that are connected in a hierarchical fashion. The first layers typically consist of a combination of convolutional filters, spatial pooling layers, and normalizations. The final few layers form an artificial neural network classifier which ultimately returns a probability (or label) for each possible class of the input (e.g., threat or non-threat). The CNN architecture is illustrated below.

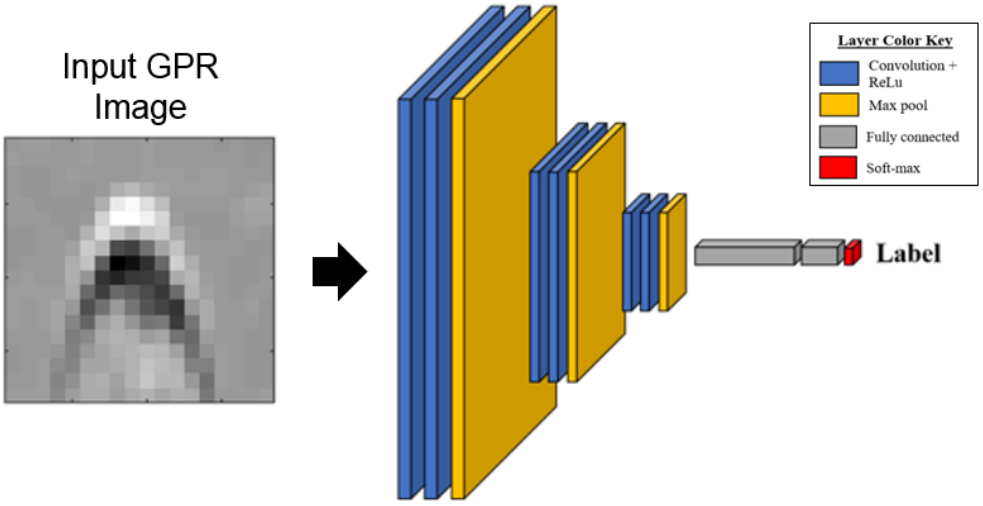

Figure 5: Illustration of the broad layout of a convolutional neural network with standard components: convolutional layers with ReLU activation units (or nonlinearities), max pooling layers, and fully connected layers (i.e., a layer that is structured like a standard neural network classifier).

CNNs, however, exhibit several characteristics that make their application to GPR buried threat detection (BTD) difficult. To achieve good performance with a CNN, an appropriate network architecture has to be specified, which requires a large number of design decisions. This problem is exacerbated by the large amounts of data that are typically needed to train the network, and the resulting long training times. For GPR in particular, obtaining sufficient amounts of data to train a network could be prohibitive. In our recent work we adapt several recent advances in the CNN literature to improve the performance of CNNs on GPR data. In particular, we investigate the following modifications to CNNs:

- We begin with a generic training procedure, where all the CNN parameters are initialized randomly, using the "Xavier-plus" initialization procedure.
- Subsequently, we use grayscale imagery from the Cifar10 dataset to pre-train our CNN.
- We then also consider using a data augmentation procedure to increase the amount of useful GPR training data.

We conduct experiments in which we sequentially add one of these techniques into the CNN training procedure. We conduct this experiment with three different CNN architectures, so that we can also explore which CNN architectural choices may yield the best results for GPR BTD. The details of the three architectures we consider are presented in Table 1.

Table 1: Description of the 3 networks employed in this work. "conv3-16" refers to a convolutional layer with 16 filters of 3×3 pixels. Maxpool(2x2, 2) refers to a pooling layer where the maximal filter response at each 2nd pixel is chosen within a window of 2×2 pixels. Finally, the fully connected layers (FC-32) represent the number of neurons (e.g., 32) in that classification layer. Note that a ReLu unit is placed after each convolutional layer (not shown here).

CNN configurations		
Input (18 × 18 image)		
(a)	(b)	(c)
Conv3-16	Conv3-16	Conv3-24
Conv3-16	Conv3-16	Conv3-24
Maxpool(2 × 2, 2)		
Conv3-32	Conv3-32	
Conv3-32	Conv3-32	
Maxpool(2 × 2, 2)		
Conv3-64		
Conv3-64		
Maxpool(2 × 2, 2)		
	FC-64	FC-32
		FC-32
FC-16		
Soft-max		

In our experiments we used a large dataset of GPR data to measure the performance of each combination of (i) CNN architecture and (ii) training modification. Cross-validation is performed to ensure that the performance reported is valid. Here we split our 8 available lanes into 4 groups, and train on 6 lanes and test on the 2 lanes being held out. This is repeated four times so that each lane is included in testing exactly once. The results from this comparison are shown in Table 2 and are reported in terms of the area under the ROC curve (AUC) up to a false alarm rate (FAR) of $0.025\ FA/m2$. A baseline for performance on this data is the previously published state-of-the-art algorithm using the HOG feature with the random forest classifier, which is shown in the final column of the table. Note that no pre-training is possible with the random forest, but the AUC when classifying the augmented dataset is shown in the final column in the table.

The results from this comparison suggest that the training procedures outlined here improve performance as they do in natural images. While initializing the network with Xavier-plus initialization achieves high detection performance, it is further benefited both by initializing the convolutional layers with pretrained layers and additionally by the dataset augmentation. Note that the HOG+RF did not improve with the augmented dataset. All CNN algorithms substantially outperform the HOG-based algorithm, although we have recently discovered that the HOG feature may not constitute a strong baseline algorithm. Future work will involve comparing the CNN algorithm to more effective conventional GPR BTD algorithms, such as those presented in the preceding section (e.g., those based on EHD or T-scan features).

Table 2: Classification performance in terms of the average AUC when testing with the trained networks from epochs 5−10 is shown for the CNN architectures defined in Table 1 along with a baseline algorithm. These algorithms are compared using the training procedures outlined in Section IV.

Training type	(a)	(b)	(c)	HOG + RF
Xavier-plus	0.9374	0.9383	0.9360	0.9254
Pre-trained	0.9398	0.9422	0.9401	--
Pre-trained + Augmentation	0.9442	0.9471	0.9451	0.9226

Feature learning methods for detecting buried threats in forward-looking ground penetrating radar (FLGPR)

Forward-looking ground penetrating radar (FLGPR) has recently been investigated as a remote sensing modality for buried target detection (e.g., landmines). In this context, raw FLGPR data is beamformed into images and then computerized algorithms are applied to automatically detect subsurface buried targets. Most existing algorithms are supervised, meaning they are trained to discriminate between labeled target and non-target imagery based on features extracted from the imagery. A large number of features have been proposed for this purpose, however thus far it has remained unclear which of them are most effective. The first part of our work provides a comprehensive comparison of detection performance using existing features on a large collection of FLGPR data. Fusion of the decisions resulting from processing each feature is also considered. The second part of our work investigated two modern feature learning approaches from the object recognition literature: the bag-of-visual-words (BOV), and the Fisher vector (FV) for FLGPR processing. These approaches have achieved state-of-the-art performance on many image recognition tasks. The FV and BOV feature learning approaches were operated on raw data and SIFT descriptors, yielding four new methods in total: BOV(Raw), BOV(SIFT), FV(Raw), and FV(SIFT). The results of comparing all the existing features, along with the proposed feature learning approaches, on all of the FLGPR polarities are shown in Figure 6. The results indicate that the new feature learning approaches generally perform well, and the best performance is achieved by the FV(SIFT). The results also indicate that overwhelmingly, the best performance was achieved on the HH polarity FLGPR images.

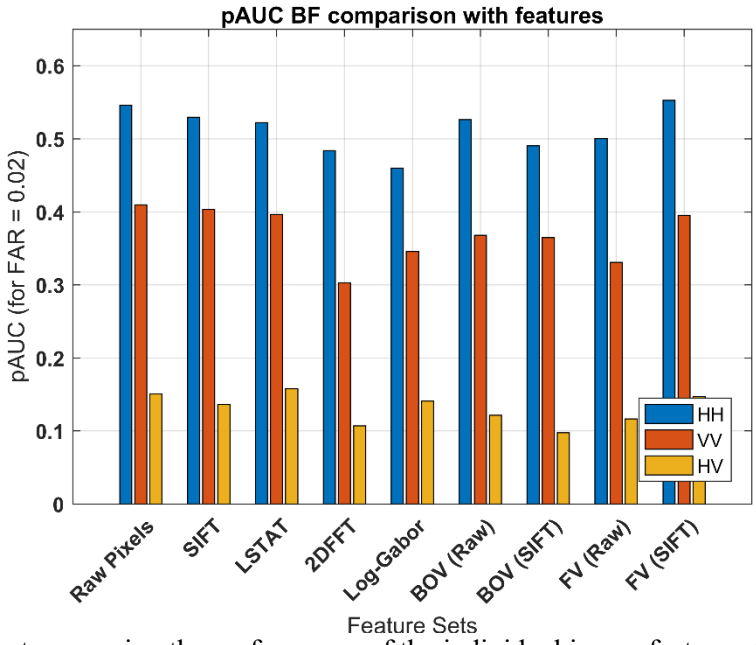

Figure 6: Bar chart comparing the performance of the individual image features. The performance measure here is the partial area under the individual ROC curve (pAUC) up to a FAR of 0.02. This measurement is normalized by the possible ROC area, meaning that pAUC = 1 was computed on an ROC curve with perfect detection at a zero false alarm rate. Here notice the HH polarization for the FLGPR outperforms the other polarizations in all features. Also notice the feature learning techniques, specifically FV do improve slightly compared to previously existing features.

In addition to comparing individual features, we also investigated decision fusion to further improve the features. Decision fusion uses an additional classifier to combine the predictions of two (or more)

classifiers in order to further improve performance. This process essentially treats the outputs of classifiers as features that are used in a subsequent classifier. Figure 7 provides a performance comparison for the prescreener, the best individual image feature set (FV(SIFT)), and the decision-level fusion result. These results indicate that further performance improvement was achieved through fusion. All algorithm training, including learning the best fusion of the features, was perfor These approaches have achieved state-of-the-art performance on many image recognition tasks. These approaches have achieved state-of-the-art performance on many image recognition tasks.

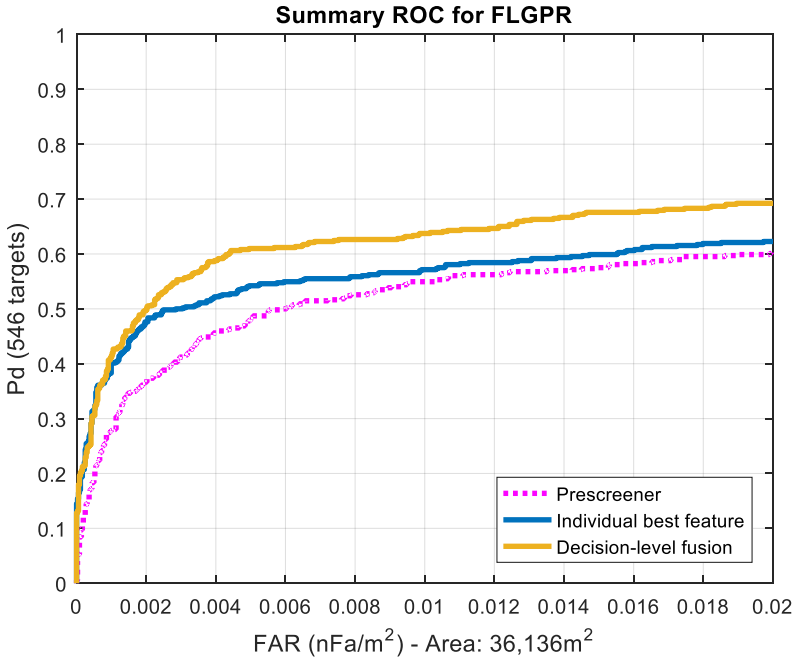

Figure 7: Receiver operating characteristic curve (ROC) for algorithms run on the FLGPR. Performance measurement with the ROC describes the probability of detection (Pd) and false alarm rate, or FAR measured as number of false alarms per meter squared. Ideal performance would have Pd = 1 at zero FAR. The three results shown here are the prescreener anomaly detector, the best single image feature set, and the decision-level fusion result.